Eye Movements: Building a Stable World from Glance to Glance

Frequent exploratory eye-movements called saccades pose for the visual system the problem of combining information from successive fixations into an apparently seamless conscious experience. A new study shows that information from successive fixations is combined, not by fusing fixation 'snapshots', but by integrating more complex visual attributes at a mid-high level of analysis.

David Burr[1,2] and Maria Concetta Morrone[3]

"[The eye] moveth it selfe with such swiftness and nimbleness, without stirring of the head, as is almost incredible" [1]. Frequent ballistic motions of the eyes called saccades form the basis of active vision, allowing us to explore and sample efficiently a large region of space with the fine-grained resolution of the tiny central retinal foveae. The challenge for the visual system is to piece together the samples from successive fixations to construct the stable representation of the world that we all consciously perceive.

Theories about trans-saccadic integration have abounded over the past decades. Early ideas (for example, see [2]) assumed the existence of a 'trans-saccadic memory buffer' which accumulates high-precision information from each saccade that is used to construct a detailed representation of the world (like pinning stamps on a tailor's dummy). These ideas fell out of favour, largely because of the implicit implication that the visual system must construct some form of stable Cartesian theatre to be viewed by a homunculus.

More recent theories have swung to the opposite extreme, assuming that perceptual stability depends, paradoxically, on a lack of internal representation of the world [3]. Several researcher have suggested that detailed visual information can readily be gleaned by making an eye movement on demand, so no visual memory is necessary [4]. In practice, however, it is still necessary for the brain to know where to look for the information that it needs, as eye movements are not random and are rarely wasted in natural tasks [5]. And there is strong evidence that information about a scene representation is accumulated across saccades [6,7]. But how?

As reported recently in *Current Biology* [8], David Melcher has examined what sorts of attributes are integrated across saccades, by clever exploitation of visual after-effects. After-effects are easy to observe: stare at a high-contrast picture for a while, and its negative will appear when you look elsewhere (see www.viperlib.com and Figure 1 for many more compelling examples). After-effects were thought to reflect adaptation of peripheral neurons, such as the photoreceptors to produce negative after-images, but more recently very intricate after-images have been described, even for images as complex as human faces [9]. Melcher [8] took advantage of the fact that after-images are long-lasting, in the order of seconds (and sometimes much longer), to investigate what sort of after-images transferred across saccades, and whether the transference was specific to spatial position ('spatiotopic').

Melcher [8] chose four types of after-effect (Figure 1), designed to tap different stages of processing: contrast adaptation (almost certainly implemented in V1 [10]); tilt and form after-effects (probably implemented at intermediate areas, such as V2 and V4 [11]); and face after-effects, involving higher

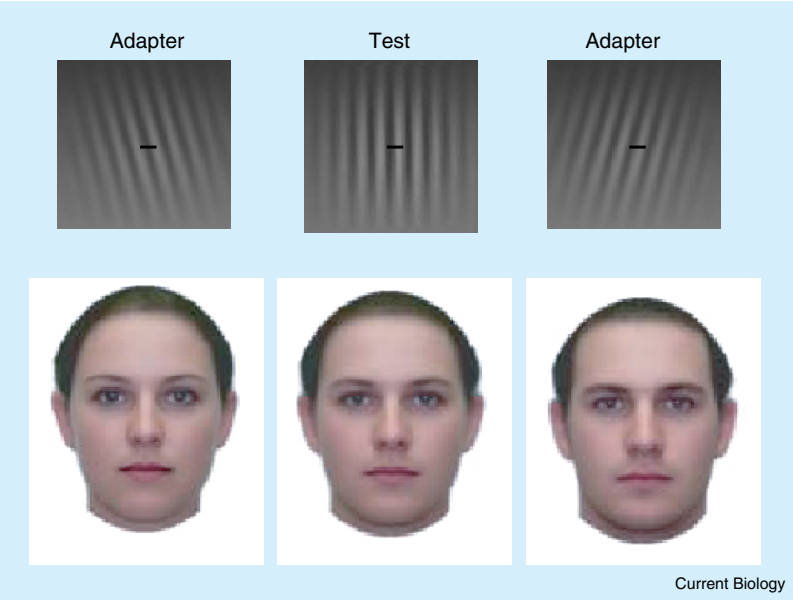

Figure 1. Illustration of two of the adaptation tasks similar to those used by Melcher [8].
The reader can experience adaptation, by first staring at one of the adapting figures (left or right) for 20 seconds or so, then observing the central test. Top: tilt after effect: after viewing the tilted adaptor (letting the eyes wander over the small bar), the vertical test grating appears tilted in the other direction. Bottom: example of a clear face after-effect (from [20]): the central face is a male–female morph; after adapting to the female face on the left, the morphed face appears male, whereas after adapting to the male on the right it appears female. In reference [8], baseline adaptation levels were first established with the eyes still, then re-measured with a saccade intervening between adaptor and test.

associative areas located in the fusiform gyrus [12]. He first established baseline strengths for the four after-effects by measuring adaptation with the eyes stationary and the adaptor and test presented to the same retinotopic (and spatiotopic) position. He then re-measured the effects, with a saccade intervening between adaptor and test, when the two stimuli were in the same spatial position (spatiotopic) and when in different spatial positions (matched in eccentricity and retinotopic displacement).

The results were beautifully clear [8]. Contrast after-effects did not transfer across saccades at all, whether in the same spatial position or not. On the other hand, face after-effects for spatially coincident (but retinally separated) stimuli transferred totally, to the same extent as for the eye stationary and retinal location matched. Interestingly, the tilt and form after-effects showed partial transfer, at about two-thirds of baseline strength. The clear implication of the study is that low-level descriptive details of images, such as local contrast, are not integrated across saccades, but high-level descriptions, such as orientation and form, and particularly human face representations, are built up over saccades. Furthermore, the effects are to a large extent spatiotopic, much stronger when the adaptor and test were in the same location in external space. These results sit well with a previous publication [13] showing that motion sequences are integrated across saccades in a spatiotopic way.

So what do these results mean? They clearly show that visual information is transferred from one fixation to another in such a way as to preserve spatiotopic representation. Basic attributes like contrast do not seem to be transferred in this way, but more elaborated representations such as faces clearly are. And most interestingly, intermediate attributes, like slope and shape, are partially transferred. It has long been known that receptive fields of visual neurons are not always fixed to their retinal locations but vary with eye position. Fully spatiotopic and craniotopic neurones have been reported in areas V6 [14] and VIP [15]. Cells in areas LIP [16], superior colliculus [17], V3A [18] and V4 [19] are at least transiently craniotopic, in that their receptive fields move in anticipation of the saccades. But this sort of behaviour has not been observed to date in area V1 or V2 [18], areas that almost certainly mediate contrast adaptation: in these areas the receptive fields remain rigidly locked to their retinal coordinates. Interestingly, in many intermediate areas like V3A, only a portion of the cells shift receptive fields pre-saccadically [18], which could explain why mid-level tasks, such as orientation and form, may be only partially affected by pre-saccadic spatiotopic adaptation.

The important conclusion from this [8] and related studies is that the visual system does combine information from one fixation to the next, but that this process is not like sticking postage stamps on a tailor's dummy: detailed 'snapshots' are not integrated within a trans-saccadic buffer that preserves the external metric [2]. Indeed such a scheme could be problematic, as scenes do change continuously as objects move and rotate: inappropriate integration could lead to a Picasso-like rendition. Trans-saccadic integration does not occur at the pixel level, but after a certain amount of visual processing, so attributes such as form, orientation, motion, and even complex entities such as faces are integrated across fixations. This in itself does not solve the problem of visual stability, but could provide a basis for visual continuity with ever-changing retinal input.

References

1. Laurentius, A. (1599). A Discourse on the Preservation of Sight: of Melancholike Diseases, and of Old Age (Oxford: OUP).
2. Jonides, J., Irwin, D.E., and Yantis, S. (1982). Integrating visual information from successive fixations. Science *215*, 192–194.
3. O'Regan, J.K., and Noe, A. (2001). A sensorimotor account of vision and visual consciousness. Behav. Brain Sci. *24*, 939–973.
4. Triesch, J., Ballard, D.H., Hayhoe, M.M., and Sullivan, B.T. (2003). What you see is what you need. J. Vis. *3*, 86–94.
5. Land, M., Mennie, N., and Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. Perception *28*, 1311–1328.
6. Melcher, D. (2001). Persistence of visual memory for scenes. Nature *412*, 401.
7. Ross, J., and Ma-Wyatt, A. (2003). Saccades actively maintain perceptual continuity. Nat. Neurosci. *7*, 65–69.
8. Melcher, D. (2005). Spatiotopic transfer of visual form adaptation across saccadic eye movements. Curr. Biol. *15*, October 11th issue.
9. Leopold, D.A., O'Toole, A.J., Vetter, T., and Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. Nat. Neurosci. *4*, 89–94.
10. Solomon, S.G., Peirce, J.W., Dhruv, N.T., and Lennie, P. (2004). Profound contrast adaptation early in the visual pathway. Neuron *42*, 155–162.
11. Gallant, J.L., Braun, J., and Van Essen, D.C. (1993). Selectivity for polar, hyperbolic, and Cartesian gratings in macaque visual cortex. Science *259*, 100–103.
12. Kanwisher, N., McDermott, J., and Chun, M.M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. J. Neurosci. *17*, 4302–4311.
13. Melcher, D., and Morrone, M.C. (2003). Spatiotopic temporal integration of visual motion across saccadic eye movements. Nat. Neurosci. *6*, 877–881.
14. Galletti, C., Battaglini, P.P., and Fattori, P. (1993). Parietal neurons encoding spatial locations in craniotopic coordinates. Exp. Brain Res. *96*, 221–229.
15. Duhamel, J., Bremmer, F., BenHamed, S., and Graf, W. (1997). Spatial invariance of visual receptive fields in parietal cortex neurons. Nature *389*, 845–848.
16. Duhamel, J.R., Colby, C.L., and Goldberg, M.E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. Science *255*, 90–92.
17. Walker, M.F., Fitzgibbon, J., and Goldberg, M.E. (1995). Neurons of the monkey superior colliculus predict the visual result of impending saccadic eye movements. J. Neurophysiol. *73*, 1988–2003.
18. Nakamura, K., and Colby, C.L. (2002). Updating of the visual representation in monkey striate and extrastriate cortex during saccades. Proc. Natl. Acad. Sci. USA *99*, 4026–4031.
19. Tolias, A.S., Moore, T., Smirnakis, S.M., Tehovnik, E.J., Siapas, A.G., and Schiller, P.H. (2001). Eye movements modulate visual receptive fields of V4 neurons. Neuron *29*, 757–767.
20. Rhodes, G., Jeffery, L., Watson, T.L., Jaquet, E., Winkler, C., and Clifford, C.W. (2004). Orientation-contingent face aftereffects and implications for face-coding mechanisms. Curr. Biol. *14*, 2119–2123.

[1]Department of Psychology, University of Florence, Italy. [2]Institute of Neuroscience of the National Research Council, Pisa, Italy. [3]Faculty of Psychology, Vita-Salute "San Raffaele" University, Milan, Italy.

DOI: 10.1016/j.cub.2005.10.003